Spatial Autoregressively Distributed Lag Models: Equivalent Forms and Estimation.

Jorgen Lauridsen

Associate Professor

Department of Statistics and Demography
University of Southern Denmark
Campusvej 55
DK-5230 Odense M
Denmark

E-mail jtl@sam.sdu.dk
Homepage http://www.ou.dk/tvf/statdem/lauridsen.html

This version was prepared for presentation at the *40th Congress of the European Regional Science Association*, 29 August to 1 September 2000, Barcelona. It represents a preliminary draft; future version is available from the author.

Abstract.

A spatial generalization of the (from times-series special case well known) Autoregressively Distributed lag model is defined. Equivalent forms - a Spatial Error Correction model, a Spatial Bewley model and a Spatial Baardsen model - are considered. As none of these may be consitently estimated by Ordinary Least Squares, an Instrument Variable estimation procedure is investigated.

1. Introduction.

Spatial regression has been discussed widely in books dedicated to developments in spatial econometrics, notably by Anselin (1988) and Anselin and Florax (1995). The consequenses for estimation and inference of the presence of stable spatial processes has been widely studied (Haining 1990, Anselin 1988, Bivand 1980, Richardson 1990, Richardson and Hèmon 1981, Clifford and Richardson 1985, Clifford, Richardson and Hèmon 1989). A recent study (Fingleton 1999) takes the first steps into analyses of implications of spatial unit roots, spatial cointegration and spatial Error Correction models.

The present paper contributes to the further development of these topics by introduction of a general Spatial Autoregressively Distributed Lag (SADL) model and different variants of this. As none of these models may be consistently estimated by Ordinary Least Squares (OLS), a consistent Instrument variable (IV) estimation procedure is investigated. The performance of this estimator is evaluated using Monte Carlo simulations while adressing the impacts of sample size and controlling for proximity structure.

2. Models for spatial dynamics.

The Spatial Autoregressive (SAR) model was initially studied by Whittle (1954) and has been used extensively in works by Ord (1975), Cliff and Ord (1981), Ripley (1981), Upton and Fingleton (1985), Anselin (1988), Griffith (1992), and Haining (1990). The SAR is defined by

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\beta + \epsilon \ , \ \epsilon \sim N(\mathbf{0} \ , \ \sigma^2\mathbf{I}) \ ,$$

in which $\mathbf{y}$ is an n×1 vector, $\mathbf{X}$ an n×K matrix of exogenous covariates, ρ the autoregression parameter, $\mathbf{I}$ the n×n identity matrix, ϵ an n×1 vector of white noises distributed with variances σ^2, and $\mathbf{W}$ an n×n proximity matrix defined by $W_{ij} = 1$ if observation j is assumed to impact observation i. $\mathbf{W}$ may be noncircular, which is the case for the times series variant where $W_{ij} = 1$ if j = i-1. For the general spatial case, $\mathbf{W}$ is generally circular. For example if the sample consists of a cross-section of n regions $\mathbf{W}$ is usually defined by $W_{ij} = W_{ji} = 1$ if region i and j are neighbours. As shown by Anselin (1988), circularity of $\mathbf{W}$ renders OLS estimation of the parametres inconsistent. This is in contrast to the times-series special case (and any other non-circular cases) where OLS provides consistent (although inefficient) estimation.

A Spatial Autoregressively Distributed Lag (SADL) model is defined by respecifying the SAR as

$$(1) \qquad \mathbf{y} = \alpha_0 + \alpha_1\mathbf{W}\mathbf{y} + \beta_0\mathbf{x} + \beta_1\mathbf{W}\mathbf{x} + \epsilon$$

where $\mathbf{x}$ and $\mathbf{y}$ are n×1 vectors, and α_0 , α_1 , β_0 , and β_1 are parametres. We may be more detailed and specify a SADL(p,q,k) defined by adding spatial lags for $\mathbf{y}$ and $\mathbf{x}$ up to order p and q, and k explanatory $\mathbf{x}$ variables. In this respect, (1) represents a SADL(1,1,1) model. However, for case of simplicity, wi will concentrate on the SADL(1,1,1) and shortly denote this SADL, as the generalization to higher order models is almost obvious: Define $\mathbf{L}$ as the spatial lag operator, i.e. $\mathbf{L}(\mathbf{x}) = \mathbf{W}\mathbf{x}$, $\mathbf{L}^2(\mathbf{x}) = \mathbf{L}(\mathbf{L}(\mathbf{x})) = \mathbf{W}(\mathbf{W}\mathbf{x}) = \mathbf{W}^2\mathbf{x}$, and $\mathbf{L}^q(\mathbf{x}) = \mathbf{L}(\ \mathbf{L}^{q-1}(\mathbf{x})) = \mathbf{W}^q\mathbf{x}$.

The SADL specifies how the expectation of y_i is formed, in terms of x_i and x_j 's and y_j 's in contiguous units. In other words, SADL is a level-to-level local specifikation. A global specification is given by unconditional expectations on the form $E(y_i) = y^*$ in (1). Using $E(\epsilon_i) = 0$, we have

$$\mathbf{y^*} = \alpha_0 + \alpha_1 \mathbf{y^*} + \beta_0 \mathbf{x^*} + \beta_1 \mathbf{x^*}$$

hence

$$\mathbf{y^*} = [\alpha_0/(1-\alpha_1)] + [(\beta_0 + \beta_1)/(1-\alpha_1)]\, \mathbf{x^*} = k_0 + k_1 \mathbf{x^*}$$

where k_1 is the global multiplier for **y** with respect to **x**, which is defined in the case of α_1 being less than 1, i.e. spatial stationarity (See Fingleton (1999) for a formal treatment of spatial (non-)stationarity).

Some easy manipulations of (1) provides the equivalent representation

(2) $$\Delta\mathbf{y} = \alpha_0 + (\alpha_1-1)\mathbf{Wy} + \beta_0\Delta\mathbf{x} + (\beta_0+\beta_1)\mathbf{Wx} + \epsilon$$

where $\Delta = (\mathbf{I}-\mathbf{W})$. Further manipulations provide

(3) $$\Delta\mathbf{y} = \alpha_0 + (\alpha_1-1)(\mathbf{Wy} - \mathbf{Wx}) + \beta_0\Delta\mathbf{x} + (\beta_0+\beta_1+\alpha_1-1)\mathbf{Wx} + \epsilon\ .$$

Alternative manipulations provide

(4) $$\mathbf{y} = \alpha_0/(1-\alpha_1) + (\alpha_1/(1-\alpha_1))\Delta\mathbf{y} + \beta_0\Delta\mathbf{x} + ((\beta_0+\beta_1)/(1-\alpha_1))\mathbf{x} - (\beta_1/(1-\alpha_1))\Delta\mathbf{x} + (1/(1-\alpha_1))\epsilon\ .$$

The forms (2)-(3)-(4) are algebraically equivalent to (1) but provide different interpretations. (2) is a spatial generalization of the times-series Baardsen specification, which we will denote the SBA model. (3) generalizes the Error Correction (EC) model and will be denoted the SEC model. Finally, (4) is a generalization of the Bewley transfom which we will call the SBE model.

Opposed to the SADL the SBA and the SEC describe the formings of expected local differences in **y** as depending on local differences in **x** and locally lagged values in **x**. They are distinctive in that the SBA introduces locally lagged levels in y whereas the SEC introduces the locally lagged discrepancy between **y** and **x**.

3. IV estimation of spatial dynamics models.

None of the specifications (1)-(4) can be estimated using OLS. This is due to the presence of contemporaneous **y** values in the variable **Wy** emerging in some form or another as an explanatory variable, implying correlation between **Wy** and ϵ. For the case of the SAR this is proved in details in Anselin (1988), whereas Fingleton (1999) provides the proof for the SEC.Their arguments are directly carried over to the SADL SBA and SBE models. Due to the aforementioned correlation asymptotically justified methodologies must be applied. Basically, two estimation methods are provided: The ML-GLS and the IV estimation.

Briefly ML-GLS consists of two steps: First the log likelihood function for **y** is concentrated to be a non-analytical function of α_1 only. Using some iterative method, the estimate of α_1

maximizing the log likelihood function is found. Second, the maximizing estimates for α_0, β_0 and β_1 are provided using one-step GLS estimators. Any sort of inference is carried out using the Fisher Information matrix. See Anselin (1988) for details and further references.

IV estimation is base on the idea of finding a variable **z** which is uncorrelated with ϵ but correlated with **Wy** (or whatever form in **y** appearing on the right-hand side of (1)-(4)) and using this as an instrument variable in a one-step least square estimation. Formally, if we want to estimate the SADL in (1), we define **X** = [**i Wy x Wx**] and **Z** = [**i z x Wx**], where **i** is an n×1 vector of 1's. Defining $\boldsymbol{\gamma}_{\mathbf{SADL}} = (\alpha_0\ \alpha_1\ \beta_0\ \beta_1)'$, the IV estimator is

$$\mathbf{g}_{\mathbf{SADL}} = (\mathbf{X'P_ZX})^{-1}\mathbf{X'P_Zy}$$

where $\mathbf{P_z} = \mathbf{Z(Z'Z)}^{-1}\mathbf{Z'}$. The covariance matrix is provided by

$$\mathbf{V}_{\mathbf{SADL}} = \sigma^2(\mathbf{X'P_zX})^{-1}$$

with σ^2 estimated consistently by

$$s^2 = (\ \mathbf{y}\text{-}\mathbf{Xg}_{\mathbf{SADL}}\ )'(\ \mathbf{y}\text{-}\mathbf{Xg}_{\mathbf{SADL}}\ )/n\ .$$

As a choice for **z**, Anselin (1988) suggests the lagged value of the prediction of **y** from an OLS regression on those variables in **X** not correlated with ϵ, i.e. **x** and **Wx**. Denoting the predicted **y** by **y^**, the instrument variable is defined as **Wy^**, and the IV estimator is obtained by setting **Z** = [**i Wy^ x Wx**].

Using **y^** for **y** in occurrences on the right-hand side, IV estimation of the alternative forms (1) - (4) is easily provided. The choices of **X**, **Z,** and dependent variable for (1)-(4) are outlined in Table 1.

Table 1. Choices of **X**, **Z**, and dependent variable.

Model	X	Z	dependent variable
(1) SADL	[**i Wy x Wx**]	[**i Wy^ x Wx**]	**y**
(2) SBA	[**i Wy Δx Wx**]	[**i Wy^ Δx Wx**]	**Δy**
(3) SEC	[**i (Wy-Wx) Δx Wx**]	[**i (Wy^-Wx) Δx Wx**]	**Δy**
(4) SBE	[**i Δy x Δx**]	[**i Δ(y^) x Δx**]	**y**

Using the one-to-one correspondence between the parametres of the four models, IV estimators for $\boldsymbol{\gamma}_{\mathbf{SADL}}$ may be derived from any of the four models upon IV estimation of these, just as the $\mathbf{V}_{\mathbf{SADL}}$ is easily derived using for example the delta method (Greene, 2000). Asymptotically, equal estimates for $\boldsymbol{\gamma}_{\mathbf{SADL}}$ and $\mathbf{V}_{\mathbf{SADL}}$ will emerge, although they may deviate for a fixed size sample. As such, the four models are asymptotically equivalent with respect to IV performance. Consequently, the success of IV in all models depends on the success of IV applied to any model. And - basically - this success hings on the success of the choice of **y^** as an instrument for occurrences of **y** in any **X** matrix. We will investigate this topic using a Monte Carlo based simulation study. Due to the asymptotic similarity of the four models, a study based on the SADL will suffice.

4. A simulation study.

The focus of our interest is the estimation of the SADL defined in (1). Two central topics must be adressed:

1. Can α_0, α_1, β_0, and β_1 be estimated consistently, using the suggested IV estimator?

2. Can meaningfull inference be derived using the estimated $\mathbf{V}_{\mathbf{SADL}}$?

Topic 2. involves inspection of the asymptotic t values for each estimated parameter, defined as

$$t = (g_p - \gamma_p)/s_p$$

where s_p is the square root of the p'th diagonal element in the estimated $\mathbf{V}_{\mathbf{SADL}}$. Further, the applicability of $\mathbf{V}_{\mathbf{SADL}}$ for model inference will be adressed by examining the Wald test for model significanse, defined as

$$\text{Wald} = (\mathbf{g}_{\mathbf{SADL}} - \boldsymbol{\gamma}_{\mathbf{SADL}})'(\mathbf{V}_{\mathbf{SADL}})^{-1}(\mathbf{g}_{\mathbf{SADL}} - \boldsymbol{\gamma}_{\mathbf{SADL}}) .$$

To ensure generality of the study, we will investigate the properties of IV for α_1 varying between 0 and 1. The resting parametres, i.e. α_0, β_0, β_1, and σ^2 will be restricted to 1, as their sign and magnitude do not provide any problems. Further the impact of varying sample size n will be investigated. Finally, to avoid restriction of the results to cases covered by any specific **W** matrix, a randomization of this matrix will be employed. This randomization is performed by the following simple rule:

For each of the n(n-1)/2 possible proximity relations: Generate a random number from the U(0,1) distribution. If this value is higher than a preselected value, d, assign $W_{ij} = W_{ji} = 1$, otherwise 0.

Full generality is obtained by repeating the study for different values of d.

The full design of the study is described in the following Monte Carlo algorithm:

For n=25, 50, 100 do:
 For d=0.01, 0.05 do:
  For α_1 = 0.01, 0.25, 0.5, 0.75, 0.9 do:
   Replicate 10,000 times:
    Generate $\boldsymbol{\epsilon}$ from n independent N(0,1)
    Generate **x** from n independent U(0,1)
    Create a random **W** using d and the above rule
    Row-standardize **W** (i.e. divide each element with rowsum)
    Calculate $\mathbf{y} = (\mathbf{I}-\mathbf{W})^{-1}(\mathbf{i} + \mathbf{x} + \mathbf{Wx} + \boldsymbol{\epsilon})$
    Perform IV estimation of SADL
    Store estimates, denoted by a_0, a_1, b_0 , and b_1
    Store t values for the parametres
    Store the Wald test
   Calculate 5, 10 50, 90, 95 per cent deciles for each stored quantity
Conclude the study by comparing these deciles to their theoretical counterparts.

The results from the Monte Carlo algorithm are collected in Table 2. Many interesting featuresmay be derived from these results. We briefly outline those of main concern for us:

The estimator a_0 is generally downward biased. This bias devreases with increasing sample size. Further, the bias seems to be larger for a high-density W matrix. The t value is also downward biased and has a strong tendency towards too short tails as compared to the N(0,1) distribution. That is, an overtendency to accept the hypothesized H0 value is present. Whereas the bias in the t value seems to decrease for increased sample size, the short-tail tendency seems to prevail. This latter prevalence is also unaffected by the density of W. In general, all these problems are worsened for increasing values of α_1.

The estimator a_1 as well as its t value is generally upward biased. This bias increases with increasing density of W, but decreases with increasing sample size. The bias increases with increasing α_1. Further, the t values have shorter tails than the N(0,1) distribution. This empirical distribution does not vary very much while α_1 n and d change.

The estimators b_0 and b_1 are remarkably stable. The - generally downward - biases are very small, even for large α_1 and are reduced when n increases. Further, the density of W does not impact the biases. The t values are generally almost unbiased, but their distributions have shorter tails tha the N(0,1).

The Wald test has a peculiar behaviour: For small sample sizes, it seems to be overstated, whereas this overstatement reduces and even turns into an understatement with increasing sample size. This behaviour seems almost unaffected by the size of α_1 and the density of W.

For empirical researchers applying the IV estimation methodology, we will suggest to account for the following features while interpreting estimation results:

- The estimate of α_1 is somewhat overstated but its t value is understated.
- The parametres for exogenous variables as well as for spatial lags of these is slighly understated as well as the corresponding t values. That is, one must not be too strict in rejecting significanse of these.
- The Wald test for model significanse is somewhat understated for fairly large sample sizes, but overstated for very small sample sizes. We suspect this feature to carry over to any asymptotic Wald-type test for model specification, based on the estimated covariance matrix (though this suspicion is not formally confirmed for any but the model significanse test).

5. An empirical illustration.

(Estimation of a commuting model for 275 Danish municipalities - provided in future version of the paper)

6. Conclusions.

(Follows in future version of the paper)

References.

Anselin, L. 1988: *Spatial Econometrics: Methods and Models*. Kluwer.
Anselin, L. and R. Florax (eds.)1995: *New Directions in Spatial Econometrics*. Springer-Verlag.
Bivand, R. 1980:A Monte-Carlo Study of Correlation Coefficient Estimation With Spatially Autocorrelated Observations. *Quaestiones Geographicae*, 6, 5-10.
Cliff, A. and J. Ord 1981: *Spatial Processes: Models and Applications*. Pion.
Clifford, P. and S. Richardson 1985: Testing the Association between Two Spatial Processes. *Statistics and Decisions*, Supplement no. 2, 155-60.
Clifford, P., S. Richardson. and D. Hémon 1989: Assessing the Significanse of the Correlation between Two Spatial Processes. *Biometrics*, 45,, 123-34.
Fingleton, B 1999: Spurious Spatial Regression: Some Monte Carlo Results with a Spatial Unit Root and Spatial Cointegration. *Journal of Regional Science*, 39, 1-19.
Greene, W. 2000: *Econometric Analysis*. Prentice-Hall.
Griffith, D. 1992: Simplifying the Normalizing Factor in Spatial Autoregressions for Irregular Lattices. *Papers in Regional Science*, 71 71-86.
Haining, R. 1990: *Spatial Data Analysis in the Social and Environmental Sciences*. Cambridge University Press.
Ord, J. 1975: Estimation Methods for Models of Spatial Interaction. *Journal of the American Statistical Association*, 70, 120-6.
Richardson, S. 1990: Some Remarks on the Testing of Association between Spatial Processes. In D. Griffith (ed.): *Spatial Statistics: Past, Present and Future*. Image. Pp. 277-309.
Richardson S. and D. Hémon 1981: On the Variance of the Sample Correlation between Two Independent Lattice Processes. *Journal of Applied Probability*, 18, 943-8.
Ripley, B. 1981: *Spatial Statistics*. Wiley.
Upton G. and B. Fingleton 1985: *Spatial Data Analysis by Example, vol. 1*. Wiley.
Whittle, P. 1954: On Stationary Processes in the Plane. *Biometrika*, 41, 434-49.

Table 2. Results of Monte Carlo study.

Deciles for theoretical distribution:	5	10	50	90	95
N(0,1):	-1.61	-1.21	0	1.21	1.61
χ2(4):	0.71	1.06	3.36	7.78	9.49

α1	n	d	dec.	Wald	a0	t(a0)	a1	t(a1)	b0	t(b0)	b1	t(b1)
0.01	25	0.01	5	0.04	-8.31	-1.47	-7.51	-1.14	-2.14	-1.24	-8.30	-1.11
			10	0.15	-3.12	-1.06	-3.55	-0.86	-0.80	-0.90	-3.67	-0.81
			50	*1.79*	*0.72*	*-0.07*	*0.02*	*0.00*	*1.01*	*0.00*	*0.98*	*0.00*
			90	6.88	6.65	0.77	3.37	0.82	2.83	0.96	4.14	1.01
			95	9.42	12.32	1.03	7.21	1.07	4.09	1.31	8.17	1.45
		0.05	5	0.05	-10.02	-1.39	-7.88	-1.20	-1.91	-1.31	-10.17	-1.17
			10	0.17	-3.88	-1.01	-3.88	-0.91	-0.72	-0.96	-4.41	-0.84
			50	*1.92*	*0.76*	*-0.05*	*-0.03*	*-0.02*	*0.99*	*-0.01*	*0.96*	*-0.01*
			90	7.06	7.10	0.86	3.97	0.82	2.68	1.00	4.63	1.00
			95	9.42	13.52	1.14	8.10	1.07	3.83	1.40	9.09	1.44
	50	0.01	5	0.06	-7.05	-1.35	-6.20	-1.09	-1.06	-1.23	-6.34	-1.00
			10	0.20	-2.65	-1.01	-3.07	-0.83	-0.24	-0.91	-2.88	-0.76
			50	*1.82*	*0.79*	*-0.07*	*0.01*	*0.00*	*0.99*	*0.00*	*1.02*	*0.00*
			90	5.77	5.64	0.75	3.09	0.83	2.26	0.91	3.30	1.02
			95	7.69	10.02	0.96	6.04	1.08	3.01	1.23	5.90	1.40
		0.05	5	0.06	-11.12	-1.31	-7.84	-1.12	-0.85	-1.13	-9.01	-1.07
			10	0.20	-4.17	-0.97	-3.83	-0.87	-0.14	-0.96	-3.82	-0.82
			50	*1.87*	*0.79*	*-0.05*	*0.00*	*0.00*	*0.98*	*-0.02*	*1.00*	*0.00*
			90	6.05	7.09	0.81	4.04	0.81	2.10	0.92	4.39	0.98
			95	7.84	12.84	1.07	8.83	1.04	2.74	1.27	8.72	1.39
	100	0.01	5	0.10	-5.50	-1.32	-4.81	-1.14	-0.30	-1.26	-4.99	-1.02
			10	0.31	-2.01	-1.01	-2.39	-0.90	0.14	-0.96	-2.21	-0.81
			50	*1.98*	*0.83*	*-0.07*	*0.03*	*0.01*	*0.99*	*-0.01*	*1.07*	*0.04*
			90	5.78	4.77	0.80	2.48	0.87	1.83	0.96	2.79	1.07
			95	7.47	8.52	1.03	5.09	1.12	2.25	1.25	4.83	1.46
		0.05	5	0.06	-13.13	-1.23	-9.43	-1.10	-0.24	-1.24	-9.36	-1.02
			10	0.20	-5.10	-0.92	-4.53	-0.85	0.22	-0.93	-4.35	-0.78
			50	*1.84*	*0.84*	*-0.04*	*-0.01*	*0.00*	*0.98*	*-0.02*	*1.11*	*0.02*
			90	5.60	8.18	0.81	4.56	0.81	1.77	0.95	4.54	0.98
			95	7.35	16.02	1.02	9.75	1.05	2.19	1.28	9.32	1.34

(table 2 continued)

α1	n	d	dec.	Wald	a0	t(a0)	a1	t(a1)	b0	t(b0)	b1	t(b1)
0.25	25	0.01	5	0.06	-11.53	-1.60	-5.89	-1.01	-2.13	-1.37	-7.66	-1.23
			10	0.18	-4.53	-1.20	-2.61	-0.75	-0.87	-1.00	-3.44	-0.92
			50	*1.96*	*0.48*	*-0.14*	*0.37*	*0.06*	*0.98*	*-0.01*	*0.88*	*-0.03*
			90	7.16	7.26	0.67	3.39	0.97	2.83	0.96	3.81	1.00
			95	9.67	14.28	0.88	6.57	1.29	4.02	1.32	6.47	1.41
		0.05	5	0.07	-14.21	-1.55	-6.56	-1.05	-1.82	-1.41	-8.72	-1.25
			10	0.20	-5.85	-1.16	-3.06	-0.89	-0.73	-1.03	-4.13	-0.93
			50	*2.05*	*0.45*	*-0.12*	*0.38*	*0.06*	*0.96*	*-0.02*	*0.92*	*-0.02*
			90	7.31	8.27	0.72	3.85	0.97	2.63	0.97	4.33	1.02
			95	9.80	15.88	0.97	7.55	1.26	3.63	1.37	7.74	1.43
	50	0.01	5	0.08	-8.16	-1.48	-4.40	-1.02	-0.88	-1.36	-5.10	-1.22
			10	0.26	-2.97	-1.17	-2.10	-0.80	-0.23	-1.03	-2.32	-0.93
			50	*2.04*	*0.69*	*-0.10*	*0.30*	*0.03*	*0.98*	*-0.02*	*0.96*	*-0.01*
			90	6.32	6.23	0.71	2.55	0.99	2.23	0.93	3.08	0.99
			95	8.32	11.24	0.91	5.15	1.28	2.99	1.25	4.68	1.35
		0.05	5	0.07	-12.94	-1.41	-7.40	-1.05	-0.82	-1.35	-7.39	-1.19
			10	0.24	-5.58	-1.07	-3.42	-0.81	-0.35	-1.02	-3.52	-0.92
			50	*1.99*	*0.58*	*-0.10*	*0.33*	*0.04*	*0.97*	*-0.02*	*0.96*	*-0.01*
			90	6.34	9.11	0.75	3.65	0.96	2.15	0.98	4.04	0.98
			95	8.34	18.12	0.96	7.21	1.22	2.75	1.28	7.57	1.37
	100	0.01	5	0.20	-4.70	-1.51	-3.14	-1.08	-1.18	-1.41	-3.15	-1.25
			10	0.46	-1.85	-1.19	-1.36	-0.87	0.19	-1.07	-1.55	-0.98
			50	*2.26*	*0.80*	*-0.08*	*0.28*	*0.03*	*0.98*	*-0.02*	*1.05*	*0.03*
			90	6.12	4.90	0.80	1.95	1.06	1.82	0.97	2.60	1.04
			95	7.85	8.47	1.00	3.37	1.33	2.24	1.27	3.48	1.38
		0.05	5	0.08	-14.75	-1.33	-8.01	-1.05	-0.14	-1.29	-7.32	-1.14
			10	0.27	-6.31	-1.03	-3.59	-0.81	0.23	-1.00	-3.62	-0.88
			50	*2.01*	*0.70*	*-0.06*	*0.30*	*0.02*	*1.00*	*0.00*	*1.06*	*0.02*
			90	5.83	9.77	0.76	4.08	0.94	1.76	0.95	4.09	0.97
			95	7.52	19.76	0.98	7.92	1.21	2.17	1.27	7.70	1.31

(table 2 continued)

α1	n	d	dec.	Wald	a0	t(a0)	a1	t(a1)	b0	t(b0)	b1	t(b1)
0.50	25	0.01	5	0.08	-13.69	-1.76	-3.64	-0.91	-1.72	-1.51	-5.55	-1.44
			10	0.24	-5.77	-1.36	-1.56	-0.67	-0.82	-1.14	-2.94	-1.10
			50	*2.17*	*0.28*	*-0.20*	*0.65*	*0.12*	*0.90*	*-0.05*	*0.84*	*-0.05*
			90	7.57	7.69	0.58	2.72	1.18	2.70	0.90	3.68	0.96
			95	9.97	15.84	0.74	4.89	1.52	3.69	1.25	5.23	1.36
		0.05	5	0.08	-17.92	-1.65	-4.75	-0.96	-1.61	-1.51	-7.37	-1.46
			10	0.25	-8.34	-1.28	-2.03	-0.72	-0.67	-1.12	-3.58	-1.08
			50	*2.24*	*0.18*	*-0.16*	*0.68*	*0.11*	*0.91*	*-0.05*	*0.83*	*-0.04*
			90	7.66	9.89	0.64	3.42	1.14	2.59	0.95	0.83	1.00
			95	10.32	19.47	0.86	6.16	1.45	3.56	1.30	2.35	1.40
	50	0.01	5	0.13	-9.00	-1.72	-2.98	-0.97	-0.73	-1.52	-3.55	-1.38
			10	0.33	-3.49	-1.33	-1.27	-0.75	-0.18	-1.14	-1.78	-1.09
			50	*2.21*	*0.57*	*-0.15*	*0.59*	*0.09*	*0.97*	*-0.02*	*0.95*	*-0.02*
			90	6.76	6.91	0.66	2.11	1.18	2.24	0.92	3.04	0.95
			95	8.79	13.22	0.86	3.68	1.50	2.90	1.22	4.03	1.26
		0.05	5	0.07	-12.94	-1.41	-7.40	-1.05	-0.82	-1.35	-7.39	-1.19
			10	0.24	-5.58	-1.07	-3.42	-0.81	-0.15	-1.02	-3.52	-0.92
			50	*1.99*	*0.58*	*-0.10*	*0.33*	*0.04*	*0.97*	*-0.01*	*0.96*	*-0.01*
			90	6.34	9.11	0.75	3.65	0.96	2.15	0.98	4.05	0.98
			95	8.34	18.12	0.97	7.22	1.22	2.75	1.28	7.57	1.37
	100	0.01	5	0.27	-4.62	-1.65	-1.95	-1.02	-0.05	-1.51	-2.09	-1.41
			10	0.56	-1.89	-1.30	-0.81	-0.82	0.25	-1.16	-0.97	-1.11
			50	*2.42*	*0.74*	*-0.11*	*0.54*	*0.05*	*0.99*	*-0.01*	*1.02*	*0.01*
			90	6.61	5.42	0.76	1.56	1.20	1.81	0.98	2.57	1.00
			95	8.28	9.36	0.94	2.44	1.50	2.17	1.27	3.17	1.32
		0.05	5	0.12	-20.95	-1.42	-5.74	-0.99	-0.11	-1.40	-6.85	-1.28
			10	0.35	-9.76	-1.11	-2.48	-0.77	0.24	-1.08	-3.39	-1.00
			50	*2.16*	*0.31*	*-0.10*	*0.66*	*0.08*	*0.97*	*-0.04*	*0.91*	*-0.03*
			90	6.12	11.63	0.73	3.84	1.06	1.72	0.94	3.59	0.95
			95	7.68	22.53	0.94	6.99	1.34	2.06	1.24	5.73	1.31

(table 2 continued)

α1	n	d	dec.	Wald	a0	t(a0)	a1	t(a1)	b0	t(b0)	b1	t(b1)
0.75	25	0.01	5	0.10	-17.02	-1.87	-1.69	-0.88	-1.34	-1.61	-3.72	-1.61
			10	0.31	-7.76	-1.49	-0.42	-0.66	-0.64	-1.24	-2.17	-1.26
			50	*2.51*	*0.06*	*-0.24*	*0.86*	*0.17*	*0.91*	*-0.05*	*0.86*	*-0.05*
			90	8.30	9.06	0.57	2.03	1.37	2.58	0.95	3.67	0.97
			95	10.98	18.86	0.78	3.23	1.73	3.37	1.30	4.85	1.32
		0.05	5	0.08	-17.93	-1.65	-4.75	-0.96	-1.61	-1.51	-7.37	-1.46
			10	0.25	-8.35	-1.27	-2.03	-0.72	-0.67	-1.12	-3.59	-1.08
			50	*2.24*	*0.18*	*-0.17*	*0.68*	*0.11*	*0.91*	*-0.05*	*0.83*	*-0.04*
			90	7.66	9.89	0.64	3.41	1.14	2.59	0.95	4.17	1.00
			95	10.23	19.47	0.87	6.16	1.44	3.56	1.30	6.30	1.40
	50	0.01	5	0.17	-11.51	-1.83	-1.29	-0.92	-0.50	-1.63	-2.16	-1.56
			10	0.43	-4.47	-1.43	-0.26	-0.74	-0.08	-1.23	-1.14	-1.25
			50	*2.47*	*0.46*	*-0.16*	*0.81*	*0.11*	*0.96*	*-0.03*	*0.96*	*-0.02*
			90	7.51	8.08	0.67	1.61	1.31	2.13	0.93	2.92	0.94
			95	9.63	15.73	0.85	2.51	1.71	2.67	1.22	3.78	1.24
		0.05	5	0.11	-29.39	-1.61	-3.13	-1.66	-0.54	-0.94	-4.60	-1.51
			10	0.32	-12.95	-1.25	-1.13	-1.26	-0.06	-0.72	-2.37	-1.14
			50	*2.33*	*-0.15*	*-0.15*	*0.88*	*-0.15*	*0.95*	*0.12*	*0.91*	*-0.04*
			90	7.00	14.90	0.68	2.74	0.68	2.04	1.21	3.50	0.97
			95	9.06	30.19	0.88	4.97	0.88	2.57	1.56	4.96	1.30
	100	0.01	5	0.32	-5.60	-1.75	-0.75	-1.04	0.05	-1.63	-1.35	-1.63
			10	0.63	-2.60	-1.38	-0.06	-0.84	0.30	-1.25	-0.61	-1.25
			50	*2.60*	*0.73*	*-0.10*	*0.77*	*0.05*	*1.00*	*0.00*	*1.00*	*0.00*
			90	7.07	6.68	0.78	1.33	1.32	1.75	1.00	2.42	1.03
			95	8.92	12.13	0.99	1.78	1.67	2.04	1.29	2.97	1.34
		0.05	5	0.16	-33.15	-1.54	-3.80	-0.98	-0.04	-1.52	-4.91	-1.43
			10	0.41	-15.22	-1.23	-1.51	-0.77	0.26	-1.16	-2.53	-1.13
			50	*2.31*	*-0.04*	*-0.11*	*0.87*	*0.09*	*0.97*	*-0.04*	*0.95*	*-0.02*
			90	6.54	17.92	0.75	3.02	1.20	1.71	0.95	3.41	0.96
			95	8.39	35.44	0.95	5.26	1.50	2.06	1.25	4.79	1.28

(table 2 continued)

α1	n	d	dec.	Wald	a0	t(a0)	a1	t(a1)	b0	t(b0)	b1	t(b1)
0.90	25	0.01	5	0.12	-29.12	-1.88	-0.35	-0.87	-1.13	-1.65	-3.11	-1.66
			10	0.36	-12.77	-1.51	0.30	-0.66	-0.53	-1.26	-1.82	-1.29
			50	*2.68*	*-0.28*	*-0.24*	*0.97*	*0.19*	*0.92*	*-0.04*	*0.87*	*-0.04*
			90	8.66	12.14	0.60	1.66	1.41	2.44	0.99	3.46	0.98
			95	11.47	24.67	0.78	2.50	1.80	3.16	1.32	4.58	1.32
		0.05	5	0.11	-53.40	-1.81	-1.48	-0.93	-1.11	-1.67	-4.48	-1.65
			10	0.35	-24.89	-1.44	-0.21	-0.69	-0.52	-1.24	-2.41	-1.25
			50	*2.65*	*-1.34*	*-0.21*	*1.02*	*0.19*	*0.92*	*-0.05*	*0.85*	*-0.05*
			90	8.53	22.07	0.65	2.27	1.40	2.42	1.02	3.89	1.03
			95	11.29	46.91	0.87	3.77	1.75	3.14	1.40	5.43	1.41
	50	0.01	5	0.19	-18.05	-1.82	-0.21	-0.92	-0.36	-1.70	-1.64	-1.67
			10	0.48	-7.60	-1.44	0.34	-0.72	-0.01	-1.29	-0.91	-1.29
			50	*2.65*	*0.26*	*-0.16*	*0.93*	*0.12*	*0.95*	*-0.04*	*0.96*	*-0.02*
			90	7.67	11.53	0.68	1.38	1.39	2.01	0.98	2.79	0.98
			95	9.85	22.40	0.86	1.91	1.76	2.45	1.29	3.50	1.28
		0.05	5	0.13	-64.00	-1.67	-1.90	-0.91	-0.50	-1.59	-3.88	-1.57
			10	0.37	-28.86	-1.31	-0.44	-0.69	-0.05	-1.20	-2.07	-1.23
			50	*2.45*	*-1.53*	*-0.16*	*1.03*	*0.15*	*0.95*	*-0.05*	*0.84*	*-0.06*
			90	7.48	26.79	0.67	2.49	1.28	1.98	0.95	3.40	0.96
			95	9.56	56.59	0.88	4.23	1.64	2.43	1.28	4.51	1.30
	100	0.01	5	0.34	-10.21	-1.79	-0.01	-1.02	0.11	-1.69	-0.98	-1.67
			10	0.63	-5.28	-1.42	0.40	-0.82	0.32	-1.29	-0.42	-1.32
			50	*2.68*	*0.57*	*-0.09*	*0.92*	*0.06*	*0.99*	*-0.01*	*0.99*	*-0.01*
			90	7.41	10.29	0.79	1.26	1.39	1.70	1.03	2.36	1.03
			95	9.34	18.27	0.98	1.53	1.75	1.96	1.34	2.83	1.33
		0.05	5	0.18	-61.02	-1.56	-3.04	-0.98	0.04	-1.60	-3.66	-1.52
			10	0.47	-32.27	-1.27	-1.00	-0.76	0.29	-1.21	-2.09	-1.18
			50	*2.49*	*-1.64*	*-0.13*	*1.03*	*0.12*	*0.98*	*-0.03*	*0.91*	*-0.04*
			90	6.77	37.57	0.75	2.66	1.24	1.72	0.98	3.43	0.98
			95	8.60	77.09	0.96	4.12	1.57	2.06	1.29	4.77	1.32

Spatial Autoregressively Distributed Lag Models: Equivalent Forms and Estimation.

Jorgen Lauridsen

Associate Professor

Department of Statistics and Demography
University of Southern Denmark
Campusvej 55
DK-5230 Odense M
Denmark

E-mail jtl@sam.sdu.dk
Homepage http://www.ou.dk/tvf/statdem/lauridsen.html

This version was prepared for presentation at the *40th Congress of the European Regional Science Association*, 29 August to 1 September 2000, Barcelona. It represents a preliminary draft; future version is available from the author.

Abstract.

A spatial generalization of the (from times-series special case well known) Autoregressively Distributed lag model is defined. Equivalent forms - a Spatial Error Correction model, a Spatial Bewley model and a Spatial Baardsen model - are considered. As none of these may be consitently estimated by Ordinary Least Squares, an Instrument Variable estimation procedure is investigated.

1. Introduction.

Spatial regression has been discussed widely in books dedicated to developments in spatial econometrics, notably by Anselin (1988) and Anselin and Florax (1995). The consequenses for estimation and inference of the presence of stable spatial processes has been widely studied (Haining 1990, Anselin 1988, Bivand 1980, Richardson 1990, Richardson and Hèmon 1981, Clifford and Richardson 1985, Clifford, Richardson and Hèmon 1989). A recent study (Fingleton 1999) takes the first steps into analyses of implications of spatial unit roots, spatial cointegration and spatial Error Correction models.

The present paper contributes to the further development of these topics by introduction of a general Spatial Autoregressively Distributed Lag (SADL) model and different variants of this. As none of these models may be consistently estimated by Ordinary Least Squares (OLS), a consistent Instrument variable (IV) estimation procedure is investigated. The performance of this estimator is evaluated using Monte Carlo simulations while adressing the impacts of sample size and controlling for proximity structure.

2. Models for spatial dynamics.

The Spatial Autoregressive (SAR) model was initially studied by Whittle (1954) and has been used extensively in works by Ord (1975), Cliff and Ord (1981), Ripley (1981), Upton and Fingleton (1985), Anselin (1988), Griffith (1992), and Haining (1990). The SAR is defined by

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\beta + \epsilon \ , \ \epsilon \sim \mathrm{N}(\mathbf{0} \ , \ \sigma^2\mathbf{I}) \ ,$$

in which $\mathbf{y}$ is an n×1 vector, $\mathbf{X}$ an n×K matrix of exogenous covariates, ρ the autoregression parameter, $\mathbf{I}$ the n×n identity matrix, ϵ an n×1 vector of white noises distributed with variances σ^2, and $\mathbf{W}$ an n×n proximity matrix defined by $W_{ij} = 1$ if observation j is assumed to impact observation i. $\mathbf{W}$ may be noncircular, which is the case for the times series variant where $W_{ij} = 1$ if $j = i-1$. For the general spatial case, $\mathbf{W}$ is generally circular. For example if the sample consists of a cross-section of n regions $\mathbf{W}$ is usually defined by $W_{ij} = W_{ji} = 1$ if region i and j are neighbours. As shown by Anselin (1988), circularity of $\mathbf{W}$ renders OLS estimation of the parametres inconsistent. This is in contrast to the times-series special case (and any other non-circular cases) where OLS provides consistent (although inefficient) estimation.

A Spatial Autoregressively Distributed Lag (SADL) model is defined by respecifying the SAR as

(1) $$\mathbf{y} = \alpha_0 + \alpha_1\mathbf{W}\mathbf{y} + \beta_0\mathbf{x} + \beta_1\mathbf{W}\mathbf{x} + \epsilon$$

where $\mathbf{x}$ and $\mathbf{y}$ are n×1 vectors, and α_0 , α_1 , β_0 , and β_1 are parametres. We may be more detailed and specify a SADL(p,q,k) defined by adding spatial lags for $\mathbf{y}$ and $\mathbf{x}$ up to order p and q, and k explanatory $\mathbf{x}$ variables. In this respect, (1) represents a SADL(1,1,1) model. However, for case of simplicity, wi will concentrate on the SADL(1,1,1) and shortly denote this SADL, as the generalization to higher order models is almost obvious: Define $\mathbf{L}$ as the spatial lag operator, i.e. $\mathbf{L}(\mathbf{x}) = \mathbf{W}\mathbf{x}$, $\mathbf{L}^2(\mathbf{x}) = \mathbf{L}(\mathbf{L}(\mathbf{x})) = \mathbf{W}(\mathbf{W}\mathbf{x}) = \mathbf{W}^2\mathbf{x}$, and $\mathbf{L}^q(\mathbf{x}) = \mathbf{L}(\ \mathbf{L}^{q-1}(\mathbf{x}))= \mathbf{W}^q\mathbf{x}$.

The SADL specifies how the expectation of y_i is formed, in terms of x_i and x_j 's and y_j 's in contiguous units. In other words, SADL is a level-to-level local specifikation. A global specification is given by unconditional expectations on the form $E(y_i) = y^*$ in (1). Using $E(\epsilon_i) = 0$, we have

$$\mathbf{y^*} = \alpha_0 + \alpha_1 \mathbf{y^*} + \beta_0 \mathbf{x^*} + \beta_1 \mathbf{x^*}$$

hence

$$\mathbf{y^*} = [\alpha_0/(1-\alpha_1)] + [(\beta_0 + \beta_1)/(1-\alpha_1)]\ \mathbf{x^*} = k_0 + k_1 \mathbf{x^*}$$

where k_1 is the global multiplier for **y** with respect to **x**, which is defined in the case of α_1 being less than 1, i.e. spatial stationarity (See Fingleton (1999) for a formal treatment of spatial (non-)stationarity).

Some easy manipulations of (1) provides the equivalent representation

(2) $$\mathbf{\Delta y} = \alpha_0 + (\alpha_1-1)\mathbf{Wy} + \beta_0\mathbf{\Delta x} + (\beta_0+\beta_1)\mathbf{Wx} + \boldsymbol{\epsilon}$$

where $\mathbf{\Delta} = (\ \mathbf{I}\text{-}\mathbf{W}\ )$. Further manipulations provide

(3) $$\mathbf{\Delta y} = \alpha_0 + (\alpha_1-1)(\ \mathbf{Wy} - \mathbf{Wx}) + \beta_0\mathbf{\Delta x} + (\beta_0+\beta_1+\alpha_1-1)\mathbf{Wx} + \boldsymbol{\epsilon}\ .$$

Alternative manipulations provide

(4) $$\mathbf{y} = \alpha_0/(1-\alpha_1) + (\alpha_1/(1-\alpha_1))\mathbf{\Delta y} + \beta_0\mathbf{\Delta x} + ((\beta_0+\beta_1)/(1-\alpha_1))\mathbf{x} - (\beta_1/(1-\alpha_1))\mathbf{\Delta x} + (1/(1-\alpha_1))\epsilon\ .$$

The forms (2)-(3)-(4) are algebraically equivalent to (1) but provide different interpretations. (2) is a spatial generalization of the times-series Baardsen specification, which we will denote the SBA model. (3) generalizes the Error Correction (EC) model and will be denoted the SEC model. Finally, (4) is a generalization of the Bewley transfom which we will call the SBE model.

Opposed to the SADL the SBA and the SEC describe the formings of expected local differences in **y** as depending on local differences in **x** and locally lagged values in **x**. They are distinctive in that the SBA introduces locally lagged levels in y whereas the SEC introduces the locally lagged discrepancy between **y** and **x**.

3. IV estimation of spatial dynamics models.

None of the specifications (1)-(4) can be estimated using OLS. This is due to the presence of contemporaneous **y** values in the variable **Wy** emerging in some form or another as an explanatory variable, implying correlation between **Wy** and $\boldsymbol{\epsilon}$. For the case of the SAR this is proved in details in Anselin (1988), whereas Fingleton (1999) provides the proof for the SEC.Their arguments are directly carried over to the SADL SBA and SBE models. Due to the aforementioned correlation asymptotically justified methodologies must be applied. Basically, two estimation methods are provided: The ML-GLS and the IV estimation.

Briefly ML-GLS consists of two steps: First the log likelihood function for **y** is concentrated to be a non-analytical function of α_1 only. Using some iterative method, the estimate of α_1

maximizing the log likelihood function is found. Second, the maximizing estimates for α_0, β_0 and β_1 are provided using one-step GLS estimators. Any sort of inference is carried out using the Fisher Information matrix. See Anselin (1988) for details and further references.

IV estimation is base on the idea of finding a variable **z** which is uncorrelated with ϵ but correlated with **Wy** (or whatever form in **y** appearing on the right-hand side of (1)-(4)) and using this as an instrument variable in a one-step least square estimation. Formally, if we want to estimate the SADL in (1), we define **X** = [**i Wy x Wx**] and **Z** = [**i z x Wx**], where **i** is an n×1 vector of 1's. Defining $\boldsymbol{\gamma}_{\mathbf{SADL}} = (\alpha_0\ \alpha_1\ \beta_0\ \beta_1)'$, the IV estimator is

$$\mathbf{g}_{\mathbf{SADL}} = (\mathbf{X'P_Z X})^{-1}\mathbf{X'P_Z y}$$

where $\mathbf{P_z} = \mathbf{Z}(\mathbf{Z'Z})^{-1}\mathbf{Z'}$. The covariance matrix is provided by

$$\mathbf{V}_{\mathbf{SADL}} = \sigma^2(\mathbf{X'P_z X})^{-1}$$

with σ^2 estimated consistently by

$$s^2 = (\ \mathbf{y}\text{-}\mathbf{X}\mathbf{g}_{\mathbf{SADL}}\ )'(\ \mathbf{y}\text{-}\mathbf{X}\mathbf{g}_{\mathbf{SADL}}\ )/n\ .$$

As a choice for **z**, Anselin (1988) suggests the lagged value of the prediction of **y** from an OLS regression on those variables in **X** not correlated with ϵ, i.e. **x** and **Wx**. Denoting the predicted **y** by **y^**, the instrument variable is defined as **Wy^**, and the IV estimator is obtained by setting **Z** = [**i Wy^ x Wx**].

Using **y^** for **y** in occurrences on the right-hand side, IV estimation of the alternative forms (1) - (4) is easily provided. The choices of **X**, **Z,** and dependent variable for (1)-(4) are outlined in Table 1.

Table 1. Choices of **X**, **Z**, and dependent variable.

Model	X	Z	dependent variable
(1) SADL	[**i Wy x Wx**]	[**i Wy^ x Wx**]	**y**
(2) SBA	[**i Wy Δx Wx**]	[**i Wy^ Δx Wx**]	**Δy**
(3) SEC	[**i (Wy-Wx) Δx Wx**]	[**i (Wy^-Wx) Δx Wx**]	**Δy**
(4) SBE	[**i Δy x Δx**]	[**i Δ(y^) x Δx**]	**y**

Using the one-to-one correspondence between the parametres of the four models, IV estimators for $\boldsymbol{\gamma}_{\mathbf{SADL}}$ may be derived from any of the four models upon IV estimation of these, just as the $\mathbf{V}_{\mathbf{SADL}}$ is easily derived using for example the delta method (Greene, 2000). Asymptotically, equal estimates for $\boldsymbol{\gamma}_{\mathbf{SADL}}$ and $\mathbf{V}_{\mathbf{SADL}}$ will emerge, although they may deviate for a fixed size sample. As such, the four models are asymptotically equivalent with respect to IV performance. Consequently, the success of IV in all models depends on the success of IV applied to any model. And - basically - this success hings on the success of the choice of **y^** as an instrument for occurrences of **y** in any **X** matrix. We will investigate this topic using a Monte Carlo based simulation study. Due to the asymptotic similarity of the four models, a study based on the SADL will suffice.

4. A simulation study.

The focus of our interest is the estimation of the SADL defined in (1). Two central topics must be adressed:

1. Can α_0, α_1, β_0, and β_1 be estimated consistently, using the suggested IV estimator?

2. Can meaningfull inference be derived using the estimated $\mathbf{V_{SADL}}$?

Topic 2. involves inspection of the asymptotic t values for each estimated parameter, defined as

$$t = (g_p - \gamma_p)/s_p$$

where s_p is the square root of the p'th diagonal element in the estimated $\mathbf{V_{SADL}}$. Further, the applicability of $\mathbf{V_{SADL}}$ for model inference will be adressed by examining the Wald test for model significanse, defined as

$$\text{Wald} = (\mathbf{g_{SADL}} - \boldsymbol{\gamma}_{\mathbf{SADL}})'(\mathbf{V_{SADL}})^{-1}(\mathbf{g_{SADL}} - \boldsymbol{\gamma}_{\mathbf{SADL}}) .$$

To ensure generality of the study, we will investigate the properties of IV for α_1 varying between 0 and 1. The resting parametres, i.e. α_0, β_0, β_1, and σ^2 will be restricted to 1, as their sign and magnitude do not provide any problems. Further the impact of varying sample size n will be investigated. Finally, to avoid restriction of the results to cases covered by any specific **W** matrix, a randomization of this matrix will be employed. This randomization is performed by the following simple rule:

For each of the n(n-1)/2 possible proximity relations: Generate a random number from the U(0,1) distribution. If this value is higher than a preselected value, d, assign $W_{ij} = W_{ji} = 1$, otherwise 0.

Full generality is obtained by repeating the study for different values of d.

The full design of the study is described in the following Monte Carlo algorithm:

For n=25, 50, 100 do:
 For d=0.01, 0.05 do:
  For α_1 = 0.01, 0.25, 0.5, 0.75, 0.9 do:
   Replicate 10,000 times:
    Generate $\boldsymbol{\epsilon}$ from n independent N(0,1)
    Generate **x** from n independent U(0,1)
    Create a random **W** using d and the above rule
    Row-standardize **W** (i.e. divide each element with rowsum)
    Calculate $\mathbf{y} = (\mathbf{I}-\mathbf{W})^{-1}(\mathbf{i} + \mathbf{x} + \mathbf{Wx} + \epsilon)$
    Perform IV estimation of SADL
    Store estimates, denoted by a_0, a_1, b_0, and b_1
    Store t values for the parametres
    Store the Wald test
   Calculate 5, 10 50, 90, 95 per cent deciles for each stored quantity
Conclude the study by comparing these deciles to their theoretical counterparts.

The results from the Monte Carlo algorithm are collected in Table 2. Many interesting featuresmay be derived from these results. We briefly outline those of main concern for us:

The estimator a_0 is generally downward biased. This bias devreases with increasing sample size. Further, the bias seems to be larger for a high-density W matrix. The t value is also downward biased and has a strong tendency towards too short tails as compared to the N(0,1) distribution. That is, an overtendency to accept the hypothesized H0 value is present. Whereas the bias in the t value seems to decrease for increased sample size, the short-tail tendency seems to prevail. This latter prevalence is also unaffected by the density of W. In general, all these problems are worsened for increasing values of α_1.

The estimator a_1 as well as its t value is generally upward biased. This bias increases with increasing density of W, but decreases with increasing sample size. The bias increases with increasing α_1. Further, the t values have shorter tails than the N(0,1) distribution. This empirical distribution does not vary very much while α_1 n and d change.

The estimators b_0 and b_1 are remarkably stable. The - generally downward - biases are very small, even for large α_1 and are reduced when n increases. Further, the density of W does not impact the biases. The t values are generally almost unbiased, but their distributions have shorter tails tha the N(0,1).

The Wald test has a peculiar behaviour: For small sample sizes, it seems to be overstated, whereas this overstatement reduces and even turns into an understatement with increasing sample size. This behaviour seems almost unaffected by the size of α_1 and the density of W.

For empirical researchers applying the IV estimation methodology, we will suggest to account for the following features while interpreting estimation results:

- The estimate of α_1 is somewhat overstated but its t value is understated.
- The parametres for exogenous variables as well as for spatial lags of these is slighly understated as well as the corresponding t values. That is, one must not be too strict in rejecting significanse of these.
- The Wald test for model significanse is somewhat understated for fairly large sample sizes, but overstated for very small sample sizes. We suspect this feature to carry over to any asymptotic Wald-type test for model specification, based on the estimated covariance matrix (though this suspicion is not formally confirmed for any but the model significanse test).

5. An empirical illustration.

(Estimation of a commuting model for 275 Danish municipalities - provided in future version of the paper)

6. Conclusions.

(Follows in future version of the paper)

References.

Anselin, L. 1988: *Spatial Econometrics: Methods and Models*. Kluwer.
Anselin, L. and R. Florax (eds.)1995: *New Directions in Spatial Econometrics*. Springer-Verlag.
Bivand, R. 1980:A Monte-Carlo Study of Correlation Coefficient Estimation With Spatially Autocorrelated Observations. *Quaestiones Geographicae*, 6, 5-10.
Cliff, A. and J. Ord 1981: *Spatial Processes: Models and Applications*. Pion.
Clifford, P. and S. Richardson 1985: Testing the Association between Two Spatial Processes. *Statistics and Decisions*, Supplement no. 2, 155-60.
Clifford, P., S. Richardson. and D. Hémon 1989: Assessing the Significanse of the Correlation between Two Spatial Processes. *Biometrics*, 45,, 123-34.
Fingleton, B 1999: Spurious Spatial Regression: Some Monte Carlo Results with a Spatial Unit Root and Spatial Cointegration. *Journal of Regional Science*, 39, 1-19.
Greene, W. 2000: *Econometric Analysis*. Prentice-Hall.
Griffith, D. 1992: Simplifying the Normalizing Factor in Spatial Autoregressions for Irregular Lattices. *Papers in Regional Science*, 71 71-86.
Haining, R. 1990: *Spatial Data Analysis in the Social and Environmental Sciences*. Cambridge University Press.
Ord, J. 1975: Estimation Methods for Models of Spatial Interaction. *Journal of the American Statistical Association*, 70, 120-6.
Richardson, S. 1990: Some Remarks on the Testing of Association between Spatial Processes. In D. Griffith (ed.): *Spatial Statistics: Past, Present and Future*. Image. Pp. 277-309.
Richardson S. and D. Hémon 1981: On the Variance of the Sample Correlation between Two Independent Lattice Processes. *Journal of Applied Probability*, 18, 943-8.
Ripley, B. 1981: *Spatial Statistics*. Wiley.
Upton G. and B. Fingleton 1985: *Spatial Data Analysis by Example, vol. 1*. Wiley.
Whittle, P. 1954: On Stationary Processes in the Plane. *Biometrika*, 41, 434-49.

Table 2. Results of Monte Carlo study.

Deciles for theoretical distribution:	5	10	50	90	95
N(0,1):	-1.61	-1.21	0	1.21	1.61
$\chi2(4)$:	0.71	1.06	3.36	7.78	9.49

α1	n	d	dec.	Wald	a0	t(a0)	a1	t(a1)	b0	t(b0)	b1	t(b1)
0.01	25	0.01	5	0.04	-8.31	-1.47	-7.51	-1.14	-2.14	-1.24	-8.30	-1.11
			10	0.15	-3.12	-1.06	-3.55	-0.86	-0.80	-0.90	-3.67	-0.81
			50	*1.79*	*0.72*	*-0.07*	*0.02*	*0.00*	*1.01*	*0.00*	*0.98*	*0.00*
			90	6.88	6.65	0.77	3.37	0.82	2.83	0.96	4.14	1.01
			95	9.42	12.32	1.03	7.21	1.07	4.09	1.31	8.17	1.45
		0.05	5	0.05	-10.02	-1.39	-7.88	-1.20	-1.91	-1.31	-10.17	-1.17
			10	0.17	-3.88	-1.01	-3.88	-0.91	-0.72	-0.96	-4.41	-0.84
			50	*1.92*	*0.76*	*-0.05*	*-0.03*	*-0.02*	*0.99*	*-0.01*	*0.96*	*-0.01*
			90	7.06	7.10	0.86	3.97	0.82	2.68	1.00	4.63	1.00
			95	9.42	13.52	1.14	8.10	1.07	3.83	1.40	9.09	1.44
	50	0.01	5	0.06	-7.05	-1.35	-6.20	-1.09	-1.06	-1.23	-6.34	-1.00
			10	0.20	-2.65	-1.01	-3.07	-0.83	-0.24	-0.91	-2.88	-0.76
			50	*1.82*	*0.79*	*-0.07*	*0.01*	*0.00*	*0.99*	*0.00*	*1.02*	*0.00*
			90	5.77	5.64	0.75	3.09	0.83	2.26	0.91	3.30	1.02
			95	7.69	10.02	0.96	6.04	1.08	3.01	1.23	5.90	1.40
		0.05	5	0.06	-11.12	-1.31	-7.84	-1.12	-0.85	-1.13	-9.01	-1.07
			10	0.20	-4.17	-0.97	-3.83	-0.87	-0.14	-0.96	-3.82	-0.82
			50	*1.87*	*0.79*	*-0.05*	*0.00*	*0.00*	*0.98*	*-0.02*	*1.00*	*0.00*
			90	6.05	7.09	0.81	4.04	0.81	2.10	0.92	4.39	0.98
			95	7.84	12.84	1.07	8.83	1.04	2.74	1.27	8.72	1.39
	100	0.01	5	0.10	-5.50	-1.32	-4.81	-1.14	-0.30	-1.26	-4.99	-1.02
			10	0.31	-2.01	-1.01	-2.39	-0.90	0.14	-0.96	-2.21	-0.81
			50	*1.98*	*0.83*	*-0.07*	*0.03*	*0.01*	*0.99*	*-0.01*	*1.07*	*0.04*
			90	5.78	4.77	0.80	2.48	0.87	1.83	0.96	2.79	1.07
			95	7.47	8.52	1.03	5.09	1.12	2.25	1.25	4.83	1.46
		0.05	5	0.06	-13.13	-1.23	-9.43	-1.10	-0.24	-1.24	-9.36	-1.02
			10	0.20	-5.10	-0.92	-4.53	-0.85	0.22	-0.93	-4.35	-0.78
			50	*1.84*	*0.84*	*-0.04*	*-0.01*	*0.00*	*0.98*	*-0.02*	*1.11*	*0.02*
			90	5.60	8.18	0.81	4.56	0.81	1.77	0.95	4.54	0.98
			95	7.35	16.02	1.02	9.75	1.05	2.19	1.28	9.32	1.34

(table 2 continued)

α1	n	d	dec.	Wald	a0	t(a0)	a1	t(a1)	b0	t(b0)	b1	t(b1)
0.25	25	0.01	5	0.06	-11.53	-1.60	-5.89	-1.01	-2.13	-1.37	-7.66	-1.23
			10	0.18	-4.53	-1.20	-2.61	-0.75	-0.87	-1.00	-3.44	-0.92
			50	*1.96*	*0.48*	*-0.14*	*0.37*	*0.06*	*0.98*	*-0.01*	*0.88*	*-0.03*
			90	7.16	7.26	0.67	3.39	0.97	2.83	0.96	3.81	1.00
			95	9.67	14.28	0.88	6.57	1.29	4.02	1.32	6.47	1.41
		0.05	5	0.07	-14.21	-1.55	-6.56	-1.05	-1.82	-1.41	-8.72	-1.25
			10	0.20	-5.85	-1.16	-3.06	-0.89	-0.73	-1.03	-4.13	-0.93
			50	*2.05*	*0.45*	*-0.12*	*0.38*	*0.06*	*0.96*	*-0.02*	*0.92*	*-0.02*
			90	7.31	8.27	0.72	3.85	0.97	2.63	0.97	4.33	1.02
			95	9.80	15.88	0.97	7.55	1.26	3.63	1.37	7.74	1.43
	50	0.01	5	0.08	-8.16	-1.48	-4.40	-1.02	-0.88	-1.36	-5.10	-1.22
			10	0.26	-2.97	-1.17	-2.10	-0.80	-0.23	-1.03	-2.32	-0.93
			50	*2.04*	*0.69*	*-0.10*	*0.30*	*0.03*	*0.98*	*-0.02*	*0.96*	*-0.01*
			90	6.32	6.23	0.71	2.55	0.99	2.23	0.93	3.08	0.99
			95	8.32	11.24	0.91	5.15	1.28	2.99	1.25	4.68	1.35
		0.05	5	0.07	-12.94	-1.41	-7.40	-1.05	-0.82	-1.35	-7.39	-1.19
			10	0.24	-5.58	-1.07	-3.42	-0.81	-0.35	-1.02	-3.52	-0.92
			50	*1.99*	*0.58*	*-0.10*	*0.33*	*0.04*	*0.97*	*-0.02*	*0.96*	*-0.01*
			90	6.34	9.11	0.75	3.65	0.96	2.15	0.98	4.04	0.98
			95	8.34	18.12	0.96	7.21	1.22	2.75	1.28	7.57	1.37
	100	0.01	5	0.20	-4.70	-1.51	-3.14	-1.08	-1.18	-1.41	-3.15	-1.25
			10	0.46	-1.85	-1.19	-1.36	-0.87	0.19	-1.07	-1.55	-0.98
			50	*2.26*	*0.80*	*-0.08*	*0.28*	*0.03*	*0.98*	*-0.02*	*1.05*	*0.03*
			90	6.12	4.90	0.80	1.95	1.06	1.82	0.97	2.60	1.04
			95	7.85	8.47	1.00	3.37	1.33	2.24	1.27	3.48	1.38
		0.05	5	0.08	-14.75	-1.33	-8.01	-1.05	-0.14	-1.29	-7.32	-1.14
			10	0.27	-6.31	-1.03	-3.59	-0.81	0.23	-1.00	-3.62	-0.88
			50	*2.01*	*0.70*	*-0.06*	*0.30*	*0.02*	*1.00*	*0.00*	*1.06*	*0.02*
			90	5.83	9.77	0.76	4.08	0.94	1.76	0.95	4.09	0.97
			95	7.52	19.76	0.98	7.92	1.21	2.17	1.27	7.70	1.31

(table 2 continued)

α_1	n	d	dec.	Wald	a0	t(a0)	a1	t(a1)	b0	t(b0)	b1	t(b1)
0.50	25	0.01	5	0.08	-13.69	-1.76	-3.64	-0.91	-1.72	-1.51	-5.55	-1.44
			10	0.24	-5.77	-1.36	-1.56	-0.67	-0.82	-1.14	-2.94	-1.10
			50	*2.17*	*0.28*	*-0.20*	*0.65*	*0.12*	*0.90*	*-0.05*	*0.84*	*-0.05*
			90	7.57	7.69	0.58	2.72	1.18	2.70	0.90	3.68	0.96
			95	9.97	15.84	0.74	4.89	1.52	3.69	1.25	5.23	1.36
		0.05	5	0.08	-17.92	-1.65	-4.75	-0.96	-1.61	-1.51	-7.37	-1.46
			10	0.25	-8.34	-1.28	-2.03	-0.72	-0.67	-1.12	-3.58	-1.08
			50	*2.24*	*0.18*	*-0.16*	*0.68*	*0.11*	*0.91*	*-0.05*	*0.83*	*-0.04*
			90	7.66	9.89	0.64	3.42	1.14	2.59	0.95	0.83	1.00
			95	10.32	19.47	0.86	6.16	1.45	3.56	1.30	2.35	1.40
	50	0.01	5	0.13	-9.00	-1.72	-2.98	-0.97	-0.73	-1.52	-3.55	-1.38
			10	0.33	-3.49	-1.33	-1.27	-0.75	-0.18	-1.14	-1.78	-1.09
			50	*2.21*	*0.57*	*-0.15*	*0.59*	*0.09*	*0.97*	*-0.02*	*0.95*	*-0.02*
			90	6.76	6.91	0.66	2.11	1.18	2.24	0.92	3.04	0.95
			95	8.79	13.22	0.86	3.68	1.50	2.90	1.22	4.03	1.26
		0.05	5	0.07	-12.94	-1.41	-7.40	-1.05	-0.82	-1.35	-7.39	-1.19
			10	0.24	-5.58	-1.07	-3.42	-0.81	-0.15	-1.02	-3.52	-0.92
			50	*1.99*	*0.58*	*-0.10*	*0.33*	*0.04*	*0.97*	*-0.01*	*0.96*	*-0.01*
			90	6.34	9.11	0.75	3.65	0.96	2.15	0.98	4.05	0.98
			95	8.34	18.12	0.97	7.22	1.22	2.75	1.28	7.57	1.37
	100	0.01	5	0.27	-4.62	-1.65	-1.95	-1.02	-0.05	-1.51	-2.09	-1.41
			10	0.56	-1.89	-1.30	-0.81	-0.82	0.25	-1.16	-0.97	-1.11
			50	*2.42*	*0.74*	*-0.11*	*0.54*	*0.05*	*0.99*	*-0.01*	*1.02*	*0.01*
			90	6.61	5.42	0.76	1.56	1.20	1.81	0.98	2.57	1.00
			95	8.28	9.36	0.94	2.44	1.50	2.17	1.27	3.17	1.32
		0.05	5	0.12	-20.95	-1.42	-5.74	-0.99	-0.11	-1.40	-6.85	-1.28
			10	0.35	-9.76	-1.11	-2.48	-0.77	0.24	-1.08	-3.39	-1.00
			50	*2.16*	*0.31*	*-0.10*	*0.66*	*0.08*	*0.97*	*-0.04*	*0.91*	*-0.03*
			90	6.12	11.63	0.73	3.84	1.06	1.72	0.94	3.59	0.95
			95	7.68	22.53	0.94	6.99	1.34	2.06	1.24	5.73	1.31

(table 2 continued)

α1	n	d	dec.	Wald	a0	t(a0)	a1	t(a1)	b0	t(b0)	b1	t(b1)
0.75	25	0.01	5	0.10	-17.02	-1.87	-1.69	-0.88	-1.34	-1.61	-3.72	-1.61
			10	0.31	-7.76	-1.49	-0.42	-0.66	-0.64	-1.24	-2.17	-1.26
			50	*2.51*	*0.06*	*-0.24*	*0.86*	*0.17*	*0.91*	*-0.05*	*0.86*	*-0.05*
			90	8.30	9.06	0.57	2.03	1.37	2.58	0.95	3.67	0.97
			95	10.98	18.86	0.78	3.23	1.73	3.37	1.30	4.85	1.32
		0.05	5	0.08	-17.93	-1.65	-4.75	-0.96	-1.61	-1.51	-7.37	-1.46
			10	0.25	-8.35	-1.27	-2.03	-0.72	-0.67	-1.12	-3.59	-1.08
			50	*2.24*	*0.18*	*-0.17*	*0.68*	*0.11*	*0.91*	*-0.05*	*0.83*	*-0.04*
			90	7.66	9.89	0.64	3.41	1.14	2.59	0.95	4.17	1.00
			95	10.23	19.47	0.87	6.16	1.44	3.56	1.30	6.30	1.40
	50	0.01	5	0.17	-11.51	-1.83	-1.29	-0.92	-0.50	-1.63	-2.16	-1.56
			10	0.43	-4.47	-1.43	-0.26	-0.74	-0.08	-1.23	-1.14	-1.25
			50	*2.47*	*0.46*	*-0.16*	*0.81*	*0.11*	*0.96*	*-0.03*	*0.96*	*-0.02*
			90	7.51	8.08	0.67	1.61	1.31	2.13	0.93	2.92	0.94
			95	9.63	15.73	0.85	2.51	1.71	2.67	1.22	3.78	1.24
		0.05	5	0.11	-29.39	-1.61	-3.13	-1.66	-0.54	-0.94	-4.60	-1.51
			10	0.32	-12.95	-1.25	-1.13	-1.26	-0.06	-0.72	-2.37	-1.14
			50	*2.33*	*-0.15*	*-0.15*	*0.88*	*-0.15*	*0.95*	*0.12*	*0.91*	*-0.04*
			90	7.00	14.90	0.68	2.74	0.68	2.04	1.21	3.50	0.97
			95	9.06	30.19	0.88	4.97	0.88	2.57	1.56	4.96	1.30
	100	0.01	5	0.32	-5.60	-1.75	-0.75	-1.04	0.05	-1.63	-1.35	-1.63
			10	0.63	-2.60	-1.38	-0.06	-0.84	0.30	-1.25	-0.61	-1.25
			50	*2.60*	*0.73*	*-0.10*	*0.77*	*0.05*	*1.00*	*0.00*	*1.00*	*0.00*
			90	7.07	6.68	0.78	1.33	1.32	1.75	1.00	2.42	1.03
			95	8.92	12.13	0.99	1.78	1.67	2.04	1.29	2.97	1.34
		0.05	5	0.16	-33.15	-1.54	-3.80	-0.98	-0.04	-1.52	-4.91	-1.43
			10	0.41	-15.22	-1.23	-1.51	-0.77	0.26	-1.16	-2.53	-1.13
			50	*2.31*	*-0.04*	*-0.11*	*0.87*	*0.09*	*0.97*	*-0.04*	*0.95*	*-0.02*
			90	6.54	17.92	0.75	3.02	1.20	1.71	0.95	3.41	0.96
			95	8.39	35.44	0.95	5.26	1.50	2.06	1.25	4.79	1.28

(table 2 continued)

$\alpha 1$	n	d	dec.	Wald	a0	t(a0)	a1	t(a1)	b0	t(b0)	b1	t(b1)
0.90	25	0.01	5	0.12	-29.12	-1.88	-0.35	-0.87	-1.13	-1.65	-3.11	-1.66
			10	0.36	-12.77	-1.51	0.30	-0.66	-0.53	-1.26	-1.82	-1.29
			50	*2.68*	*-0.28*	*-0.24*	*0.97*	*0.19*	*0.92*	*-0.04*	*0.87*	*-0.04*
			90	8.66	12.14	0.60	1.66	1.41	2.44	0.99	3.46	0.98
			95	11.47	24.67	0.78	2.50	1.80	3.16	1.32	4.58	1.32
		0.05	5	0.11	-53.40	-1.81	-1.48	-0.93	-1.11	-1.67	-4.48	-1.65
			10	0.35	-24.89	-1.44	-0.21	-0.69	-0.52	-1.24	-2.41	-1.25
			50	*2.65*	*-1.34*	*-0.21*	*1.02*	*0.19*	*0.92*	*-0.05*	*0.85*	*-0.05*
			90	8.53	22.07	0.65	2.27	1.40	2.42	1.02	3.89	1.03
			95	11.29	46.91	0.87	3.77	1.75	3.14	1.40	5.43	1.41
	50	0.01	5	0.19	-18.05	-1.82	-0.21	-0.92	-0.36	-1.70	-1.64	-1.67
			10	0.48	-7.60	-1.44	0.34	-0.72	-0.01	-1.29	-0.91	-1.29
			50	*2.65*	*0.26*	*-0.16*	*0.93*	*0.12*	*0.95*	*-0.04*	*0.96*	*-0.02*
			90	7.67	11.53	0.68	1.38	1.39	2.01	0.98	2.79	0.98
			95	9.85	22.40	0.86	1.91	1.76	2.45	1.29	3.50	1.28
		0.05	5	0.13	-64.00	-1.67	-1.90	-0.91	-0.50	-1.59	-3.88	-1.57
			10	0.37	-28.86	-1.31	-0.44	-0.69	-0.05	-1.20	-2.07	-1.23
			50	*2.45*	*-1.53*	*-0.16*	*1.03*	*0.15*	*0.95*	*-0.05*	*0.84*	*-0.06*
			90	7.48	26.79	0.67	2.49	1.28	1.98	0.95	3.40	0.96
			95	9.56	56.59	0.88	4.23	1.64	2.43	1.28	4.51	1.30
	100	0.01	5	0.34	-10.21	-1.79	-0.01	-1.02	0.11	-1.69	-0.98	-1.67
			10	0.63	-5.28	-1.42	0.40	-0.82	0.32	-1.29	-0.42	-1.32
			50	*2.68*	*0.57*	*-0.09*	*0.92*	*0.06*	*0.99*	*-0.01*	*0.99*	*-0.01*
			90	7.41	10.29	0.79	1.26	1.39	1.70	1.03	2.36	1.03
			95	9.34	18.27	0.98	1.53	1.75	1.96	1.34	2.83	1.33
		0.05	5	0.18	-61.02	-1.56	-3.04	-0.98	0.04	-1.60	-3.66	-1.52
			10	0.47	-32.27	-1.27	-1.00	-0.76	0.29	-1.21	-2.09	-1.18
			50	*2.49*	*-1.64*	*-0.13*	*1.03*	*0.12*	*0.98*	*-0.03*	*0.91*	*-0.04*
			90	6.77	37.57	0.75	2.66	1.24	1.72	0.98	3.43	0.98
			95	8.60	77.09	0.96	4.12	1.57	2.06	1.29	4.77	1.32